Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 99 (2008) 1698–1716
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Exact distribution of the generalized Wilks's statistic and applications

T. Pham-Gia

Universite de Moncton, Departement de mathematiques et de statistique, Moncton, Canada E1A 3E9

Received 31 July 2006
Available online 12 February 2008

Abstract

We establish the exact expression of the density of Wilks's statistic $\Lambda(n, p, q)$, and also those of the densities of the product and ratio of two independent such statistics, in terms of Meijer functions, and provide applications with numerical illustrations in various domains of Multivariate Analysis.

AMS subject classifications: 62H10; 62H15

Keywords: Matrix variate Beta distribution; Wilks's statistic; Product; Ratio; Meijer function; Linear discriminant; Regression

1. Introduction

Wilks's statistic $\Lambda(n, p, q)$, also sometimes denoted by $U_{n.p,q}$, is widely used for various statistical tests in multivariate analysis since it, supposedly, plays the same role as the Fisher–Snedecor F_{ν_1,ν_2} in univariate statistics. The distribution of Wilks's statistic is, however, difficult to track, because, so far, its density lacks a closed form expression, except for some simple values of its parameters [1,16]. For the same reason, derived distributions, such as the densities of the product and ratio of two independent Wilks's statistics, cannot be computed, and one has to resort to numerical methods. Various approaches have been suggested in the literature to approximate the density of $\Lambda(n, p, q)$. For example, the use of a multiple of a χ^2 variable as an approximation to $\log_e \Lambda$ is quite good, but the precision of the derived densities, based on these approximations, is often difficult to be determined, although in some recent works, some

E-mail address: thu.pham-gia@umoncton.ca.

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2008.01.021

bounds for the associated error have been established [19,5]. In the same spirit, Ulyanov, Wakaki and Fujikoshi [18] give the Berry–Esseen bound for high dimensional asymptotic approximation of $\Lambda(n, p, q)$.

It was established that $\Lambda(n, p, q)$ has the same density as a product of independent univariate beta variables, and this is the approach, which has been only partially exploited in the past, that we will use in this article. Consul [4], for example, used Meijer functions to express that density in a theoretical context, when simple values of the parameters are considered. But computation difficulties associated with the **G**-function at that time have prevented further development in that direction, and as a result, no precise numerical value of the density of $\Lambda(n, p, q)$, and of its cdf, can be computed. In this article, we will first establish the density of the general $\Lambda(n, p, q)$ under a closed form expression, as a **H**-function, or as a **G**-function probability density, resulting in the possibility of computing and drawing these densities using standard mathematical packages, such as MAPLE or MATHEMATICA. Percentiles of this distribution can now be obtained simply and accurately, and this exact distribution can now efficiently serve as a reference, to which all approximate distributions can be compared. Considering two independent generalized Wilks's statistics $\Lambda_1(n_1, p_1, q_1)$ and $\Lambda_2(n_2, p_2, q_2)$, where the parameters can now take non-integer values too, we will establish expressions of the densities of $P = \Lambda_1\Lambda_2$, and $R = \Lambda_1/\Lambda_2$, also in closed form. The likelihood ratio statistic, intimately related to Wilks's statistic, will also be considered, in different contexts.

Wilks's statistic is often compared to the Fisher–Snedecor variable F_{ν_1,ν_2} in univariate statistics but it should rather be compared to the standard beta variable there, and to the matrix variate beta, in matrix variate analysis. The determinant of the matrix beta variate, of both types I and II, will be studied here, and an alternate form of the generalized Wilks's statistic, much more closely related to F_{ν_1,ν_2}, will be presented. In the applications section, we will provide evidence for the usefulness of these exact densities.

In Section 2 we recall the results related to the basic operations on Meijer and Fox function random variables, and establish the expressions of the densities of P and R, expressed with these functions. Section 3 presents the case of the beta distribution while Section 4 deals with Wilks's Lambda and its product and quotient. Section 5 discusses other variables intimately related to the lambda and, finally, Section 6 presents different contexts in multivariate statistics, where P and R are encountered, and gives their densities in these specific cases. Several numerical examples in the same section clearly show that the densities of Wilks's statistics can now be computed, and used, in spite of their complex theoretical expressions.

2. Meijer and Fox function random variables

2.1

Definition 1. The Meijer function $\mathbf{G}(x)$, and the Fox function $\mathbf{H}(x)$, are defined as follows: $\mathbf{G}(x) = \mathbf{G}^{m\,r}_{p\,q}\left[x\left|\begin{smallmatrix}a_1,\ldots,a_p\\ b_1,\ldots,b_q\end{smallmatrix}\right.\right]$ is the integral along the complex contour L, i.e.

$$\mathbf{G}^{m\,r}_{p\,q}\left[x\left|\begin{matrix}a_1,\ldots,a_p\\ b_1,\ldots,b_p\end{matrix}\right.\right] = \frac{1}{2\pi\mathrm{i}}\int_L \frac{\prod_{j=1}^{m}\Gamma\left(b_j - s\right)\prod_{j=1}^{r}\Gamma\left(1 - a_j + s\right)}{\prod_{j=m+1}^{q}\Gamma\left(1 - b_j + s\right)\prod_{j=r+1}^{p}\Gamma\left(a_j - s\right)} x^s \mathrm{d}s.$$

It is a special case, when $\alpha_i = \beta_j = 1, \forall i, j$, of Fox's **H**-function, defined as:

$$\mathbf{H}^{m\,r}_{p\,q}\left[x \left| \begin{matrix}(a_1,\alpha_1),\ldots,(a_p,\alpha_p)\\(b_1,\beta_1),\ldots,(b_q,\beta_q)\end{matrix}\right.\right] = \frac{1}{2\pi\mathrm{i}}\int_L \frac{\prod_{j=1}^{m}\Gamma\left(b_j-\beta_j s\right)\prod_{j=1}^{r}\Gamma\left(1-a_j+\alpha_j s\right)}{\prod_{j=m+1}^{q}\Gamma\left(1-b_j+\beta_j s\right)\prod_{j=r+1}^{p}\Gamma\left(a_j-\alpha_j s\right)} x^s \mathrm{d}s.$$

Under some fairly general conditions on the poles of the gamma functions in the numerator, the above integrals exist. For discussions on the **G**-function see [9], and on the **H**-function, see [6]. Springer [17] treats some uses of these functions in Statistics.

Several common positive random variables can be expressed as **H**-function random variables since they have a density of the form:

$$f(x) = k\mathbf{H}^{m\,r}_{p\,q}\left[cx \left| \begin{matrix}(a_1,\alpha_1),\ldots,(a_p,\alpha_p)\\(b_1,\beta_1),\ldots,(b_q,\beta_q)\end{matrix}\right.\right] \quad x>0,$$

with moments $\mu_r = \mathbf{M}_{r+1}\{\mathbf{H}^{mr}_{pq}[cx|\ {}^{(a_1,\alpha_1),\ldots,(a_p,\alpha_p)}_{(b_1,\beta_1),\ldots,(b_q,\beta_q)}]\}$, where $\mathbf{M}_s$ is the Mellin transform, defined by $\mathbf{M}_s\{f(x)\} = \int_0^\infty x^{s-1}f(x)\mathrm{d}x$, for $f(x)$ defined on R^+, with its inverse Mellin transform $f(x) = \frac{1}{2\mathrm{i}\pi}\int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} x^{-s}\mathbf{M}_s(f(x))\mathrm{d}s$, valid under quite general conditions. We then have:

$$\mu_k = \frac{k}{c^{r+1}}\frac{\prod_{j=1}^{m}\Gamma\left(b_j+\beta_j+\beta_j r\right)\prod_{j=1}^{r}\Gamma\left(1-a_j-\alpha_j-\alpha_j r\right)}{\prod_{j=m+1}^{q}\Gamma\left(1-b_j-\beta_j-\beta_j r\right)\prod_{j=r+1}^{p}\Gamma\left(a_j+\alpha_j+\alpha_j r\right)}.$$

For example, the beta variable, defined on (0, 1), with density $f(x) = x^{\gamma-1}(1-x)^{\delta-1}/B(\gamma,\delta), 0 \le x \le 1, \gamma, \delta > 0$, which can be expressed as:

$$\frac{\Gamma(\gamma+\delta)}{\Gamma(\gamma)}\mathbf{H}^{1\,0}_{1\,1}\left[x\left|\begin{matrix}\gamma+\delta-1,1\\\gamma-1,1\end{matrix}\right.\right], \quad x>0, \tag{1}$$

is, in fact, a **G**-function random variable, since both the second parameters are 1. Hence,

$$f(x) = \frac{\Gamma(\gamma+\delta)}{\Gamma(\gamma)}\mathbf{G}^{1\,0}_{1\,1}\left[x\left|\begin{matrix}\gamma+\delta-1\\\gamma-1\end{matrix}\right.\right]. \tag{2}$$

2.2. *Product and ratio of* **G**-*function random variables*

We present below the expressions of the parameters related to these two basic operations [17]. These expressions, given under a convenient form easy to apply, will be used to establish subsequent results.

(a) *Product:* Let $X_j, j = 1,\ldots,s$, be s independent **H**-function random variables, each with pdf:

$$f_j(x_j) = k_j\mathbf{H}^{m_j\,r_j}_{p_j\,q_j}\left[c_j x_j \left| \begin{matrix}(a_{j,1},\alpha_{j,1}),\ldots, & ,(a_{j,p_j},\alpha_{j,p_j})\\(b_{j,1},\beta_{j,1}),\ldots, & ,(b_{j,q_j},\beta_{j,q_j})\end{matrix}\right.\right], \quad x_j>0.$$

The product of these variables, $P = \prod_{j=1}^{s} X_j$, is also a **H**-Function Random variable, with density:

$$f_P(y) = \left(\prod_{j=1}^{s} k_j\right) \mathbf{H}_{P\ Q}^{M\ R}\left[\left(\prod_{j=1}^{s} c_j\right) y \,\middle|\, \begin{matrix} (u.s.)\ (a_{1,1}, \alpha_{1,1}), \ldots, & , (a_{s,p_s}, \alpha_{s,p_s}) \\ (l.s.)\ (b_{1,1}, \beta_{1,1}), \ldots, & , (b_{s,q_s}, \beta_{s,q_s}) \end{matrix}\right],$$
$$y > 0, \tag{3}$$

where $M = \sum_{j=1}^{s} m_j$, $R = \sum_{j=1}^{s} r_j$, $P = \sum_{j=1}^{s} p_j$, $Q = \sum_{j=1}^{s} q_j$, and the two-parameter sequences, (u.s.) and (l.s.), are as follows:

(*) The upper sequence of parameters (u.s.), of total length P, consists of s consecutive subsequences of the type $(a_{1,1}, \alpha_{1,1}), \ldots, (a_{1,r_1}, \alpha_{1,r_1})$, followed by s consecutive subsequences of the type $(a_{1,(r_1+1)}, \alpha_{1,(r_1+1)}), \ldots, (a_{1,p_1}, \alpha_{1,p_1})$, i.e. we have:

$$\begin{aligned}&(a_{1,1}, \alpha_{1,1}), \ldots, (a_{1,r_1}, \alpha_{1,r_1}); (a_{2,1}, \alpha_{2,1}), \ldots, (a_{2,r_2}, \alpha_{2,r_2}); \ldots; (a_{s,1}, \alpha_{s,1}), \ldots, \\ &\quad (a_{s,r_s}, \alpha_{s,r_s}) :: (a_{1,(r_1+1)}, \alpha_{1,(r_1+1)}), \ldots, (a_{1,p_1}, \alpha_{1,p_1}); (a_{2,(r_2+1)}, \alpha_{2,(r_2+1)}), \ldots, \\ &\quad (a_{2,p_2}, \alpha_{2,p_2}) \ldots; (a_{s,(r_s+1)}, \alpha_{s,(r_s+1)}), \ldots, (a_{s,p_s}, \alpha_{s,p_s}).\end{aligned} \tag{4}$$

(**) Similarly, the lower parameter sequence (l.s.), of total length Q, consists of s consecutive subsequences of the type $(b_{1,1}, \beta_{1,1}), \ldots, (b_{1,m_1}, \beta_{1,m_1})$, followed by s consecutive subsequences of the type: $(b_{1,(m_1+1)}, \beta_{1,(m_1+1)}), \ldots, (b_{1,q_1}, \beta_{1,q_1})$, i.e. we have:

$$\begin{aligned}&(b_{1,1}, \beta_{1,1}), \ldots, (b_{1,m_1}, \beta_{1,m_1}); (b_{2,1}, \beta_{2,1}), \ldots, (b_{2,m_2}, \beta_{2,m_2}); \ldots; (b_{s,1}, \beta_{s,1}), \ldots, \\ &\quad (b_{s,m_s}, \beta_{s,m_s}) :: (b_{1,(m_1+1)}, \beta_{1,(m_1+1)}), \ldots, (b_{1,q_1}, \beta_{1,q_1}); (b_{2,(m_2+1)}, \beta_{2,(m_2+1)}), \ldots, \\ &\quad (b_{2,q_2}, \beta_{2,q_2}) \ldots; (b_{s,(m_s+1)}, \beta_{s,(m_s+1)}), \ldots, (b_{s,q_s}, \beta_{s,q_s}).\end{aligned} \tag{5}$$

(b) *Ratio*: For the ratio $W = X_1/X_2$, its density is:

$$\begin{aligned}f_W(w) &= A\mathbf{H}_{p_1+q_2\ q_1+p_2}^{m_1+r_2\ m_2+r_1} \\ &\times \left[\frac{c_1}{c_2} w \,\middle|\, \begin{matrix} (u.s.)\ (a_{1,1}, \alpha_{1,1}), \ldots, & , \ldots, (1 - b_{2,q_2} - 2\beta_{2,q_2}, \beta_{2,q_2}) \\ (l.s.)\ (b_{1,1}, \beta_{1,1}), \ldots, & , \ldots, (1 - a_{2,p_2} - 2\alpha_{2,p_2}, \alpha_{2,p_2}) \end{matrix}\right], \quad w > 0\end{aligned} \tag{6}$$

where $A = (\frac{k_1 k_2}{c_2^2})$.

(*) The upper sequence of parameters (u.s.), of total length $p_1 + q_2$, consists of the four consecutive subsequences:

$$\left.\begin{aligned}&(a_{11}, \alpha_{11}), \ldots (a_{1r_1}, \alpha_{1r_1}) \quad \text{of length } r_1 \\ &(1 - b_{2,1} - 2\beta_{2,1}, \beta_{2,1}), \ldots, (1 - b_{2,m_2} - 2\beta_{2,m_2}, \beta_{2,m_2}) \quad \text{of length } m_2 \\ &(a_{1,r_1+1}, \alpha_{1,r_1+1}), \ldots, (a_{1,p_1}, \alpha_{1,p_1}) \quad \text{of length } p_1 - r_1, \text{ and} \\ &(1 - b_{2,m_2+1} - 2\beta_{2,m_2+1}, \beta_{2,,m_2+1}), \ldots, (1 - b_{2,q_2} - 2\beta_{2,q_2}, \beta_{2,q_2}) \quad \text{of length } q_2 - m_2.\end{aligned}\right. \tag{7}$$

(**) The lower sequence (l.s.), of total length $p_2 + q_1$, also has 4 subsequences of respective lengths $m_1, r_2, q_1 - m_1$ and $p_2 - r_2$:

$$\begin{aligned}&(b_{1,1}, \beta_{1,1}), \ldots, (b_{1,m_1}, \beta_{1,m_1}) \\ &(1 - a_{2,1} - 2\alpha_{2,1}, \alpha_{2,1}), \ldots, (1 - a_{2,r_2} - 2\alpha_{2,r_2}, \alpha_{2,r_2}), \\ &(b_{1,m_1+1}, \beta_{1,m_1+1}), \ldots, (b_{1,q_1}, \beta_{1,q_1}), \quad \text{and} \\ &(1 - a_{2,r_2+1} - 2\alpha_{2,r_2+1}, \alpha_{2,r_2+1}), \ldots, (1 - a_{2,p_2} - 2\alpha_{2,p_2}, \alpha_{2,p_2}).\end{aligned} \tag{8}$$

Proof. The proofs of the above results, based on the Mellin transform of a function $f(x)$ defined on R^+, and its inverse Mellin transform, as defined previously, are quite involved, but can be found in [17, p. 214]. $\square$

3. Case of the beta variable

As seen above, the standard beta variable can have its density expressed as a **G**-function by (2).

3.1. Product

The Product X of n independent betas X_i, i.e. $X_i \sim \text{beta}(\gamma_i, \delta_i)$, $i = 1, \ldots, n$, has its density defined on (0, 1), expressed as a Meijer function as follows:

$$h(x) = \begin{cases} \left(\prod_{j=1}^{n} \dfrac{\Gamma(\gamma_i+\delta_i)}{\Gamma(\gamma_i)}\right) \mathbf{G}_{n\,n}^{n\,0}\left[x \middle| \begin{matrix} \gamma_1+\delta_1-1, \ldots, \gamma_n+\delta_n-1 \\ \gamma_1-1, \ldots, \gamma_n-1 \end{matrix}\right], & x > 0. \\ 0, \quad x \le 0. \end{cases} \tag{9}$$

It is then immediate that the product $X_1 X_2$, where $X_1$and X_2 are both products of n_1 and n_2 independent betas respectively, can be computed, using (3):

$$h(x) = \begin{cases} K\mathbf{G}_{n\,n}^{n\,\mathbf{0}}\left[x \middle| \begin{matrix} \gamma_{1,1}+\delta_{1,1}-1, \ldots, \gamma_{1,n_1}+\delta_{1,n_1}-1, \gamma_{2,1}+\delta_{2,1}-1, \ldots, \gamma_{2,n_2}+\delta_{2,n_2}-1 \\ \gamma_{1,1}-1, \ldots, \gamma_{1,n_1}-1, \gamma_{2,1}-1, \ldots, \gamma_{2,n_2}-1 \end{matrix}\right], \\ x > 0 \\ 0, \quad x \le 0, \end{cases} \tag{10}$$

with $n = n_1 + n_2$ and $K = \prod_{i=1}^{2}(\prod_{j=1}^{n_i} \frac{\Gamma(\gamma_{ij}+\delta_{ij})}{\Gamma(\gamma_{ij})})$.

Generalizing to s variables, we have:

Theorem 1. *The product of s independent variables X_j, $j = 1, \ldots, s$, with each of them being the product of n_j independent beta variables, is also a variable defined on* (0, 1)*, with density of the same type as* (10)*. We have:*

$$f(x) = \begin{cases} K\mathbf{G}_{n^*\,n^*}^{n^*\,\mathbf{0}}\left[x \middle| \begin{matrix} \gamma_{1,1}+\delta_{1,1}-1, \ldots, \gamma_{1,n_1}+\delta_{1,n_1}-1, \ldots, \gamma_{s,1}+\delta_{s,1}-1, \ldots, \gamma_{s,n_s}+\delta_{s,n_s}-1 \\ \gamma_{1,1}-1, \ldots, \gamma_{1,n_1}-1, \ldots, \gamma_{s,1}-1, \ldots, \gamma_{s,n_s}-1 \end{matrix}\right], \\ x > 0 \\ 0, \quad x \le 0 \end{cases} \tag{11}$$

with $n^* = \sum_{i=1}^{s} n_i$ and $K = \prod_{i=1}^{s}(\prod_{j=1}^{n_i} \frac{\Gamma(\gamma_{ij}+\delta_{ij})}{\Gamma(\gamma_{ij})})$.

Proof. The proof is immediate, using the relations established above. $\square$

Remark. 1. For the general case, an expression of the above densities, using only common functions, is possible. Springer [17, p. 105] gives (9) as a function of x^a and $\log x$, but the whole expression remains highly complex. Other similar expressions, equally complex, have been obtained by Mathai [10].

3.2. Ratio

When X_1 and X_2 are independent products of n_1 and n_2 beta variables respectively, as given by (9), taking into consideration the values of $r_1 = r_2 = 0, , m_1 = p_1 = q_1 = n_1$ and $m_2 = p_2 = q_2 = n_2$, we can see that the first and last subsequences in the upper parameter sequence (u.s.) of (7) are nil. For the lower parameter sequence (l.s.), the two middle subsequences of (8) are nil. The ratio then has as density:

$$f_W(w) = A\mathbf{H}_{n_1+n_2}^{n_1}{}_{n_1+n_2}^{n_2} \times \left[\frac{c_1}{c_2} w \middle| \begin{matrix} (1-b_{2,1}-2\beta_{2,1}, \beta_{2,1}), \ldots, (1-b_{2,n_2}-2\beta_{2,n_2}, \beta_{2,n_2}), (a_{1,1}, \alpha_{1,1}), \ldots, (a_{1,n_1}, \alpha_{1,n_1}) \\ (b_{1,1}, \beta_{1,1}), \ldots, (b_{1,n_1}, \beta_{1,n_1}), (1-a_{2,1}-2\alpha_{2,1}, \alpha_{2,1}), \ldots, , (1-a_{2,n_2}-2\alpha_{2,n_2}, \alpha_{2,n_2}) \end{matrix} \right],$$

$w > 0$, where for $A = (\frac{K}{c_2^2}) = \frac{1}{c_2^2} \prod_{i=1}^{2} (\prod_{j=1}^{n_i} \frac{\Gamma(\gamma_j+\delta_j)}{\Gamma(\gamma_j)})$, $c_1 = c_2 = 1$, and $\forall i, j$, we have: $a_{i,j} = \gamma_{i,j} + \delta_{i,j} - 1$, $b_{i,j} = \gamma_{i,j} - 1$. Since $\alpha_{i,j} = \beta_{i,j} = 1$, we have, in fact, a **G**-function, and hence, we have:

Theorem 2. *Let X_1 and X_2 be independent products of n_1 and n_2 beta variables respectively. Then $W = X_1/X_2$ has as density*

$$\begin{aligned} f_W(w) = {} & A\mathbf{G}_{n_1+n_2}^{n_1}{}_{n_1+n_2}^{n_2} \\ & \times \left[w \middle| \begin{matrix} -\gamma_{2,1}, \ldots, -\gamma_{2,n_2}; \gamma_{1,1}+\delta_{1,1}-1, \ldots, \lambda_{1,n_1}+\delta_{1,n_1}-1 \\ \gamma_{1,1}-1, , \ldots, \gamma_{1,n_1}-1; -\gamma_{2,1}-\delta_{2,1}, \ldots, , -\gamma_{2,n_2}-\delta_{2,n_2} \end{matrix} \right], \\ & w > 0. \end{aligned} \tag{12}$$

4. Distribution of Wilks's lambda

4.1

Since there are some differences in the definition of Wilks's criterion in the existing literature, we will follow [8] for its definition: Let $\mathbf{A} \sim W_p(m_\mathbf{A}, \mathbf{\Sigma})$ and $\mathbf{B} \sim W_p(m_\mathbf{B}, \mathbf{\Sigma})$, with $\mathbf{A}$ and $\mathbf{B}$ independent, be $(p \times p)$ Wishart matrices, and $m_A, m_B > p$. Then Wilks's statistic, or criterion, denoted by $\Lambda(n, p, q)$, where $q = m_B$ and $n = m_A + m_B$, is the ratio of two determinants, i.e. $\Lambda(n, p, q) = \frac{|\mathbf{A}|}{|\mathbf{A}+\mathbf{B}|} = \prod_{i=1}^{p} \frac{1}{1+\lambda_i}$, where λ_i are the latent roots of $|\mathbf{B}|/|\mathbf{A}|$. $\Lambda(n, p, q)$, with n, p, q positive integers, is encountered in several domains of multivariate analysis. But $\mathbf{B}$ is central Wishart, i.e. $\mathbf{B} \sim W_p(m_\mathbf{B}, \mathbf{\Sigma})$ only under some null hypothesis H_0. Otherwise, $\mathbf{B}$ has a non-central Wishart distribution, i.e. $\mathbf{B} \sim W_p(m_\mathbf{B}, \mathbf{\Sigma}, \Omega)$, and Wilks's statistic becomes non-central, i.e. $\Lambda \sim \Lambda_{nc}(n, p, q, \Omega)$.

We always suppose $m_A > p$ (or $n > p + m_B$), but for $q = m_B < p$, $\mathbf{B}$ is the sum of independent square lengths of m_B normal variables $Z_i \sim N_p(\boldsymbol{\mu}, \mathbf{\Sigma})$, i.e. $\mathbf{B} = \sum_{i=1}^{m_B} \mathbf{Z}_i' \mathbf{Z}_i$, with $\boldsymbol{\mu} = \mathbf{0}$ under H_0. The following result is well-known (see [8]):

If $q \geq p$, $\Lambda(n, p, q)$ is the product of p beta variables, i.e. it has the same distribution as

$$\prod_{i=1}^{p} T_i, \quad \text{with } T_i \sim \text{beta}((n-q-(i-1))/2, q/2). \tag{13}$$

If $q < p$, it suffices to interchange p with q in (10), i.e. we have

$$\prod_{i=1}^{q} T_i^*, \quad \text{with } T_i^* \sim \text{beta}((n-p-(i-1))/2, p/2).$$

Moreover, we have $\Lambda(n, p, q) = \Lambda(n, q, p)$, i.e. p and q are interchangeable (for another type of notation, for example, $\Lambda(p, \nu_B, \nu_A)$, we have [15, p. 120]): $\Lambda(p, \nu_B, \nu_A) = \Lambda(\nu_B, p, \nu_A + \nu_B - p)$. Hence, Wilks's statistic $\Lambda(n, p, q)$, under the null hypothesis, is a positive, unidimensional random variable, defined on (0, 1), with integral valued parameters $n, p, q \geq 1$, where $n \geq \max(p, q)$.

4.2. Expression of the density of $\Lambda(n, p, q)$

By (13) we have $\gamma_i = \frac{n-q+1-i}{2}$, $\delta_i = \frac{q}{2}$, $i = 1, p$ in (10), and hence:

Theorem 3. *For integers n, p and q, with $n \geq q \geq p$, the density of $\Lambda(n, p, q)$, is*

$$h(x) = \begin{cases} K\mathbf{G}^{p\,0}_{p\,p}\left[x\left|\begin{matrix}\frac{n}{2}-1, \frac{n-1}{2}-1\ldots, \frac{n-(p-1)}{2}-1\\ \frac{n-q}{2}-1, \frac{n-q-1}{2}-1\ldots, \frac{n-q-(p-1)}{2}-1\end{matrix}\right.\right], & x > 0\\ 0, & x < 0, \end{cases} \tag{14}$$

where $K = \prod_{i=1}^{p} \frac{\Gamma(\frac{n+1-i}{2})}{\Gamma(\frac{n-q+1-i}{2})}$.

Expression (14) is, in fact, valid for any positive values of (α, β, γ), with $\alpha \geq \beta \geq \gamma$. The related variable $\Lambda(\alpha, \beta, \gamma)$ is then called generalized Wilks's statistic. As an example, Fig. 1 gives the graph of the density of $\Lambda(12.5, 4.16, 8.34)$.

Remarks. (1) Looking from the standpoint of a product of p beta variables, $\Lambda(n, p, q)$ has several obvious properties that are traditionally presented as special cases. For example, $\Lambda(n, 1, q)$ is the beta$(\frac{n-q}{2}, \frac{q}{2})$ variable, and the relation $\frac{n-q}{q}\frac{1-\Lambda}{\Lambda} \sim F_{q,n-q}$ merely reflects the change from a beta variable to the corresponding F-variable [13]. For $p = 2$, we have the product of two independent betas, and it can be shown that $\frac{n-q-1}{q}\frac{1-\Lambda^{1/2}}{\Lambda^{1/2}} \sim F_{2q,2(n-q-1)}$.

For $q = 1$, and any value of p, by interchanging p and q, we have similar results (see [8, p. 300]).

(2) Depending on the parity of q, we can further characterize the univariate beta distributions of the product that are equivalent to $\Lambda(n, p, 2r)$ or $\Lambda(n, p, 2r + 1)$. Also, for p even or n even, $\Lambda(n, p, q)$ can be shown to be a combination of terms $-(\log u)^k u^l$, with $k =$ integer and $l =$ half-integer [1, p. 306].

(3) For small values of the parameters, Anderson [1, p. 304], Mathai and Rathie [11], among others, have derived the expression of the above density, using only elementary functions (see the section on the Likelihood ratio statistic). Schatzoff [16] was the first to compute the distribution function of $\Lambda(n, p, q)$ in closed form, using convolutions, when p and q are even.

4.3. Percentiles of Wilks's statistic's distribution

It is usual practice that Wilks's statistic be approximated by a Chi-square variable, according to Bartlett's formula: $-(n - \frac{p+q+1}{2}) \log_e \Lambda(n, p, q) \sim \chi^2_{pq}$. Hence, for the product P and ratio R, $\log_e P$ and $\log_e R$ can be approximated by a sum or difference of scaled Chi-square variables, respectively. However, it is difficult to measure the goodness of this approach, which is more accurate when n is large. As seen above, for small values of p, a transformation of Λ will lead

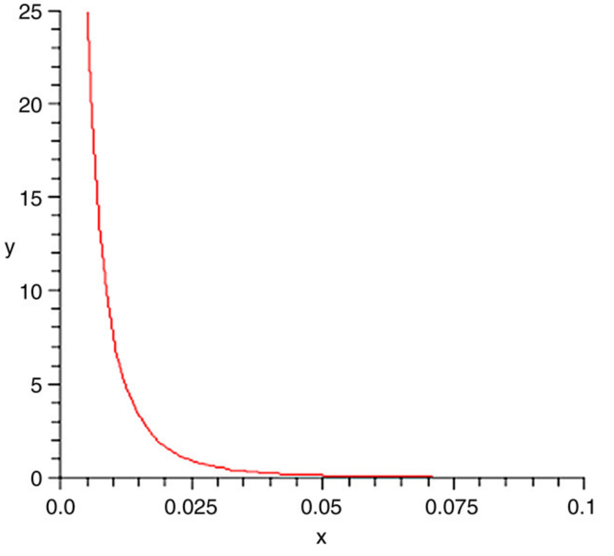

Fig. 1. Density of $\Lambda(12.5, 4.16, 8.34)$.

Table 1
Percentiles of $\Lambda(n, p, q)$ for $\alpha = 0.05$

n	$q = \mathbf{2}$		**n**	$q = \mathbf{11}$	
	Wall's value	Our value		Wall's value	Our value
18	0.111	0.1111383382	**27**	0.000987	0.0009467435
26	0.272	0.2717245961	**35**	0.010	0.01043369085
62	0.627	0.6272783556	**71**	0.161	0.1608954665
102	0.761	0.7613295146	**111**	0.333	0.3301031606

to the Fisher–Snedecor distribution and operations of the F_{ν_1,ν_2} distribution are studied in [13]. Rao and Box also provide their approximation approaches, but in general, for values of n, p, q satisfying the conditions of none of the above approaches, the approximate values obtained are less accurate.

With the closed form expression (14) given above (for $p \leq q$, with an interchange of p and q in the other case), we can now accurately determine the values of the percentiles, for any value of the trio (n, p, q). Table 1 gives, for example, values of $\Lambda_{0.05,(n,p,q)}$ for some values of (n, p, q), and compares them to those given by Wall (see [15, p. 427–434]). As we can see, the corresponding values are very close, with ours providing some extra precision. Also, the same percentiles are given for some more extreme values of p and q, but special integration techniques are required here to prevent overflow. A computer program using MAPLE can be available upon request, with special instructions to handle extreme values of p and q.

4.4. Product

Theorem 4. *Let $\Lambda_1(n_1, p_1, q_1)$ and $\Lambda_2(n_2, p_2, q_2)$ be two generalized independent Wilks's statistics. Then, for $p_1 \leq q_1$, and $p_2 \leq q_2$, $P = \Lambda_1(n_1, p_1, q_1)\Lambda_2(n_2, p_2, q_2)$ has density,*

$$h(x) = K\mathbf{G}^{p_1+p_2\ \mathbf{0}}_{p_1+p_2\ p_1+p_2} \times \left[x \,\middle|\, \begin{matrix} \frac{n_1}{2}-1, \frac{n_1-1}{2}-1, \ldots, \frac{n_1-(p_1-1)}{2}-1, \frac{n_2}{2}-1, \frac{n_2-1}{2}-1, .., \frac{n_2-(p_2-1)}{2}-1 \\ \frac{n_1-q_1}{2}-1, \frac{n_1-q_1-1}{2}-1, \ldots, \frac{n_1-q_1-(p_1-1)}{2}-1, \frac{n_2-q_2}{2}-1, \frac{n_2-q_2-1}{2}-1, \ldots, \frac{n_2-q_2-(p_2-1)}{2}-1 \end{matrix} \right] \tag{15}$$

for $x > 0$, *with* $K = \prod_{j=1}^{2} \prod_{i=1}^{p_j} \frac{\Gamma(\frac{n_j+1-i}{2})}{\Gamma(\frac{n_j-q_j+1-i}{2})}$.

Remarks. (1) For $p_1 \leq q_1$, $p_2 \geq q_2$, as well as in other situations, we have a similar result, after interchanging p_2 and q_2.

(2) An immediate consequence of the above formula is the following relation, obtained by setting $\frac{n_1-(p_1-1)}{2} - 1 = \frac{n_2}{2} - \frac{1}{2}$:

$$\Lambda(n + p_1, p_1, q)\Lambda(n, p_2, q) = \Lambda(n + p_1, p_1 + p_2, q). \tag{16}$$

This equation, which is essentially related to the number of variables, is used in several contexts in the published literature, as shown in the following sections.

4.5. Ratio

Similarly, we have:

Theorem 5. *Let* $\Lambda_1(n_1, p_1, q_1)$ *and* $\Lambda_2(n_2, p_2, q_2)$ *be two independent generalized Wilks's statistics. Depending on the relative values of* p_1 *and* q_1*, and of* p_2 *and* q_2*, their ratio* $R = \Lambda_1(n_1, p_1, q_1)/\Lambda_2(n_2, p_2, q_2)$ *has its density under a form similar to the one below. For example, when* $p_1 \leq q_1$ *and* $p_2 \leq q_2$*,we have:*

$$f(u) = K\mathbf{G}^{p_1\ \ p_2}_{p_1+p_2\ p_1+p_2} \times \left(u \,\middle|\, \begin{matrix} -\frac{n_2-q_2}{2}, -\frac{n_2-q_2-1}{2} \ldots, -\frac{n_2-q_2-(p_2-1)}{2}, \frac{n_1}{2}-1, \frac{n_1-1}{2}-1 \ldots, \frac{n_1-(p_1-1)}{2}-1 \\ \frac{n_1-q_1}{2}-1, \frac{n_1-q_1-1}{2}-1 \ldots, \frac{n_1-q_1-(p_1-1)}{2}-1, -\frac{n_2}{2}, -\frac{n_2-1}{2} \ldots, \frac{n_2-(p_2-1)}{2} \end{matrix} \right), \tag{17}$$

$0 < u < \infty$, *where* $K = \prod_{j=1}^{2} \prod_{i=1}^{p_j} \frac{\Gamma(\frac{n_j+1-i}{2})}{\Gamma(\frac{n_j-q_j+1-i}{2})}$.

Proofs. We apply (10) for the product and (11) for the ratio, using the above values of $\gamma_{i,j}$ and $\delta_{i,j}$. □

Remarks. 1. Similar formulas exist for the cases $p_1 \leq q_1$, where $p_2 > q_2$, and $p_1 > q_1$, with either $p_2 \leq q_2$ or $p_2 > q_2$.

2. We can see that although the product of two independent generalized Wilks's statistics is, in general, another Wilks's statistic, it is so for their ratio only in some situations where in the expression of the integrand of the **H**-function, there are simplifications in the Gamma functions in the numerator and the denominator. In the general case, (17) gives the precise expression of the density of the ratio of two independent generalized Wilks's statistic, that could be required in several situations.

It should also be noted here that division is NOT the inverse of multiplication for random variables, due to the dependence between the product and its factors. For example, (16) gives $\Lambda(n+p_1, p_1+p_2, q)$ as a product of two Wilks's statistics, which is also defined on (0, 1). However, in (16) the ratio $\Lambda(n+p_1, p_1+p_2, q)/\Lambda(n+p_1, p_1, q)$, has its density defined on $(0,\infty)$ by (17), and cannot be identical to $\Lambda(n, p_2, q)$. This point has caused some confusion in the literature. For this reason, if $\mathbf{T}=\mathbf{XY}$, we define $\mathbf{Y}$ as the *pseudo-ratio* $\mathbf{T}/\mathbf{X}$, denoted by $\mathbf{Y}=pr\frac{\mathbf{T}}{\mathbf{X}}$, and similarly for $\mathbf{X}$. Hence, $\Lambda(n, p_2, q)$ is the pseudo-ratio of the above two variables.

5. Variables related to Wilks's statistic

5.1. Distribution of the likelihood ratio statistic

The likelihood ratio statistic W is, in general, a function of one or several Wilks's statistics. Several authors (e.g. [1, p. 192]), have given the closed form expression of the density of $\Lambda = W^{n/2}$, under small values of the parameters, as already mentioned. For $W=\frac{|\mathbf{A}|}{\prod_{i=1}^{q}|\mathbf{A}_{ii}|}$, Consul [4] has used Meijer functions for these expressions, under a variety of cases. This is also carried out by Anderson [1, p. 236] and Mathai [10]. Our formula (19), using $\mathbf{H}$-functions, is valid for all cases.

We have: $W=\Lambda(n,p,q)^{\zeta}$, with $\zeta>0$, in the simplest case. The density of W can be obtained as follows:

First, let $f(x)=k\mathbf{H}_{p\,q}^{m\,n}\left[cx\,\middle|\,\begin{matrix}(a_1,\alpha_1),\ldots,(a_p,\alpha_p)\\(b_1,\beta_1),\ldots,(b_q,\beta_q)\end{matrix}\right]$, $x>0$ be the density of X.

Then $Y=X^{\zeta}$ has as density

$$f(y)=kc^{\zeta-1}\mathbf{H}_{p\,q}^{m\,n}\left[c^{\zeta}y\,\middle|\,\begin{matrix}(a_1+\alpha_1(1-\zeta),\alpha_1\zeta),\ldots,(a_p+\alpha_p(1-\zeta),\alpha_p\zeta)\\(b_1+\beta_1(1-\zeta),\beta_1\zeta),\ldots,(b_q+\beta_q(1-\zeta),\beta_q\zeta)\end{matrix}\right],\quad x>0.$$

If X is the product of n independent betas, using (1) and (3), X^{ζ} has as density:

$$h(x)=\begin{cases}\left(\prod_{j=1}^{n}\frac{\Gamma(\gamma_i+\delta_i)}{\Gamma(\gamma_i)}\right)\mathbf{H}_{n\,n}^{n\,0}\left[x\,\middle|\,\begin{matrix}(\gamma_1+\delta_1-\zeta,\zeta)\ldots,(\gamma_n+\delta_n-\zeta,\zeta)\\(\gamma_1-\zeta,\zeta),\ldots,(\gamma_n-\zeta,\zeta)\end{matrix}\right], & x>0\\ 0, & x\le 0.\end{cases}\tag{18}$$

Setting $\gamma_i=\frac{n_1-q_1+1-i}{2}$, $\delta_i=\frac{q_1}{2}$, $i=1,\ldots,p_1$, in (18), and the density of $W=[\Lambda_1(n_1,p_1,q_1)]^{\zeta}$, under null hypothesis, is then

$$h(x)=\begin{cases}B\mathbf{H}_{p_1\,p_1}^{p_1\,0}\left[w\,\middle|\,\begin{matrix}(\frac{n_1}{2}-\zeta,\zeta),(\frac{n_1-1}{2}-\zeta,\zeta),..,(\frac{n_1-(p_1-1)}{2}-\zeta,\zeta)\\(\frac{n_1-q_1}{2}-\zeta,\zeta),(\frac{n_1-q_1-1}{2}-\zeta,\zeta),\ldots,(\frac{n_1-q_1-(p_1-1)}{2}-\zeta,\zeta)\end{matrix}\right], & x>0\\ 0, & x\le 0,\end{cases}\tag{19}$$

where $B=\prod_{j=1}^{n_1}\frac{\Gamma(\frac{n_1-(i-1)}{2})}{\Gamma(\frac{n_1-q_1-(i-1)}{2})}$, with the value of ζ is to be determined according to the test.

5.2. Determinant of the matrix variate beta

The positive definite symmetric random matrix $\mathbf{X}$ has a beta distribution of the first kind, denoted by $\mathbf{X}\sim\text{beta}_p^I(a,b)$, if its density is of the form:

$$f(\mathbf{X}) = \frac{|\mathbf{X}|^{a-\frac{1}{2}(p+1)} \left|\mathbf{I}_p - \mathbf{X}\right|^{b-\frac{1}{2}(p+1)}}{\boldsymbol{\beta}_p(a,b)}, \quad \mathbf{0} < \mathbf{X} < \mathbf{I}_p, \tag{20}$$

where $a > \frac{1}{2}(p+1)$, $b > \frac{1}{2}(p+1)$ are positive real numbers, and $\boldsymbol{\beta}_{\mathbf{p}}(a,b)$ is the beta function in R^p, i.e. $\boldsymbol{\beta}_p(a,b) = \frac{\Gamma_p(a)\Gamma_p(b)}{\Gamma_p(a+b)}$, with $\Gamma_p(a) = \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^p \Gamma(a - \frac{i-1}{2})$.

We consider $\mathbf{U} = \det(\mathbf{X})$. Then $\mathbf{U}$ has a generalized Wilks's distribution $\Lambda(2(a+b), p, 2b)$. Hence, its density has the same form as the product of p independent univariate beta distributions, $\prod_{j=1}^p X_j$, with independent $X_j \sim \text{beta}(a - \frac{j-1}{2}, b)$. The matrix variate beta distribution arises in the familiar context:

Let $\mathbf{A} \sim W_p(m_{\mathbf{A}}, \boldsymbol{\Sigma})$ and $\mathbf{B} \sim W_p(m_{\mathbf{B}}, \boldsymbol{\Sigma})$, with $\mathbf{A}$ and $\mathbf{B}$ independent $(p \times p)$ positive definite symmetric Wishart matrices, and $m_A, m_B > p$. Then the ratio $\mathbf{U} = (\mathbf{A}+\mathbf{B})^{-1/2}\mathbf{A}(\mathbf{A}+\mathbf{B})^{-1/2}$ has the matrix beta distribution $\text{beta}_p^I(m_A/2, m_B/2)$, while $|\mathbf{U}| = \Lambda(m_A + m_B, p, m_B)$.

Concerning the product and ratio of two independent matrix beta variates, we have:

Theorem 6. *If* $\mathbf{X}_1 \sim \text{beta}_p^I(a_1, b_1)$ *is independent of* $\mathbf{X}_2 \sim \text{beta}_p^I(a_2, b_2)$*, then their product* $\mathbf{U} = (\mathbf{X}_2)^{1/2}\mathbf{X}_1(\mathbf{X}_2)^{1/2}$ *and ratio* $\mathbf{V} = (\mathbf{X}_2)^{-1/2}\mathbf{X}_1(\mathbf{X}_2)^{-1/2}$ *have their densities given by Bekker, Roux and Pham-Gia* [3]. *For their determinants,*

(a) $|\mathbf{U}|$ *has density:*

$$f(u) = K\mathbf{G}_{2p,\,2p}^{2p,\,0}\left[u\left|\begin{matrix} a_1+b_1-1, a_1+b_1-\frac{3}{2}, \ldots, a_1+b_1-\frac{1}{2}(p+1), a_2+b_2-1, a_2+b_2-\frac{3}{2}, \ldots, a_2+b_2-\frac{1}{2}(p+1) \\ a_1-1, a_1-\frac{3}{2}\ldots, a_1-\frac{1}{2}(p+1), a_2-1, a_2-\frac{3}{2}, \ldots, a_2-\frac{1}{2}(p+1) \end{matrix}\right.\right],$$
$$0 < u < 1,$$

where $K = \prod_{i=1}^2 \prod_{j=1}^p \frac{\Gamma(a_i+b_i+\frac{j-1}{2})}{\Gamma(a_i+\frac{j-1}{2})}$*, and*

(b) $|\mathbf{V}|$ *has density:*

$$f(v) = K\mathbf{G}_{2p\;2p}^{p,\,p}\left[v\left|\begin{matrix} -a_2, -a_2+\frac{1}{2}, \ldots, -a_2+\frac{1}{2}(p-1), a_1+b_1-1, a_1+b_1-\frac{3}{2}, \ldots, a_1+b_1-\frac{1}{2}(p+1) \\ a_1-1, a_1-\frac{3}{2}\ldots, a_1-\frac{1}{2}(p+1), -a_2-b_2, -a_2-b_2+\frac{1}{2}, \ldots, -a_2-b_2+\frac{1}{2}(p-1) \end{matrix}\right.\right],$$
$$0 < u < \infty.$$

Proof. It suffices to apply Theorems 4 and 5 to the product and ratio of $\Lambda(2(a_1+b_1), p, 2b_1)$ and $\Lambda(2(a_2+b_2), p, 2b_2)$. $\square$

Remark. We would arrive at the same result by remarking that $|\mathbf{U}|$ is the product of two products of independent betas, i.e. $|\mathbf{X}_i| = \prod_{j=1}^p W_j$, with $W_j \sim \text{beta}(a_i - \frac{j-1}{2}, b_i)$, $i = 1, 2$. Applying Theorem 1, we have the above expression of $|\mathbf{U}|$, and similarly for $|\mathbf{V}|$.

5.3. Alternate form of Wilks's statistic

Similarly, we define the positive definite symmetric random matrix $\mathbf{Y}$ to have a beta of the second kind distribution, denoted by $\mathbf{Y} \sim \text{beta}_p^{II}(a, b)$, if its density is of the form:

$$f(\mathbf{Y}) = \frac{|\mathbf{Y}|^{a-\frac{1}{2}(p+1)}}{\boldsymbol{\beta}_p(a,b)\left|\mathbf{I}_p + \mathbf{Y}\right|^{a+b}}, \quad \mathbf{0} < \mathbf{Y}, \tag{21}$$

where $a > \frac{1}{2}(p+1)$, $b > \frac{1}{2}(p+1)$ are positive real numbers.

We know that the two types of beta distributions can be transformed into each other by simple transformations. Moreover, $|\mathbf{Y}|$ can be expressed as a product of p independent univariate betaprimes as follows:

$|\mathbf{Y}| = \prod_{j=1}^{p} T_j, 0 \leq |\mathbf{Y}| < \infty$,with $T_j \sim$ betaprime$(a - \frac{(j-1)}{2}, b - \frac{(j-1)}{2})$, where $X \sim$ betaprime(α, β) if its univariate density has the form:

$$f(x) = \frac{1}{B(\alpha,\beta)} \frac{x^{\alpha-1}}{(1+x)^{\alpha+\beta}} = \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \mathbf{H}_{1\,1}^{1\,1}\left[\begin{matrix}(-\beta,1)\\(\alpha-1,1)\end{matrix}\right].$$

Expressed as a **G**-function density, using (3), the density of $|\mathbf{Y}|$ is:

$$f(x) = A.\mathbf{G}_{p\,p}^{p\,p}\left[x \left| \begin{matrix} -b, -\left(b-\frac{1}{2}\right), \ldots, -\left(b - \frac{p-1}{2}\right) \\ a-1, a-\frac{3}{2}, \ldots, a-\frac{p+1}{2} \end{matrix}\right.\right],$$

with

$$A = \prod_{j=1}^{p} \frac{1}{\Gamma(a - \frac{(j-1)}{2})\Gamma(b - \frac{(j-1)}{2})}.$$

Similarly to the case of $\Lambda(n,p,q)$, let us consider $\mathbf{A} \sim W_p(m_\mathbf{A}, \mathbf{\Sigma})$ and $\mathbf{B} \sim W_p(m_\mathbf{B}, \mathbf{\Sigma})$, with $\mathbf{A}$ and $\mathbf{B}$ independent $(p \times p)$ positive definite symmetric Wishart matrices, and $m_A, m_B > p$. Then the ratio $\mathbf{V} = \mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2}$ has a beta$_p^{II}(m_A/2, m_B/2)$ distribution if $\mathbf{\Sigma} = \gamma\mathbf{I}_p, \gamma > 0$, and $\mathbf{B}^{1/2}$ is symmetric. Then $|\mathbf{V}| \sim \Lambda'((m_A + m_B), p, m_B)$, and hence $|\mathbf{V}| = \prod_{j=1}^{p}$ betaprime$(\frac{m_A}{2} - \frac{(j-1)}{2}, \frac{m_B}{2} - \frac{(j-1)}{2})$, $0 \leq |\mathbf{V}| < \infty$. For $p = 1$, the Fisher–Snedecor variable F_{ν_1,ν_2} obtained is just a multiple of the univariate beta prime Y.

As remarked by Kshirsagar [8, p. 292], Wilks's ratio is defined as $\frac{|\mathbf{A}|}{|\mathbf{A}+\mathbf{B}|}$ for tractability and its relation to the likelihood ratio criterion. But if we define $\Lambda'(n,p,q) = \frac{|\mathbf{A}|}{|\mathbf{B}|}$, this statistic would be more similar to the univariate F_{ν_1,ν_2}. The following definition is based on (6).

Definition 2. For integral values of n, p, q, with $n \geq q \geq p$, Wilks's statistic of the second kind, $\Lambda'(n,p,q)$, has density

$$f(x) = A\mathbf{G}_{p\,p}^{p\,p}\left[x \left| \begin{matrix} -\frac{q}{2}, -\left(\frac{q}{2}-\frac{1}{2}\right), \ldots, -\left(\frac{q}{2} - \frac{p-1}{2}\right) \\ \frac{n-q}{2}-1, \frac{n-q}{2}-\frac{3}{2}, \ldots, \frac{n-q}{2}-\frac{p+1}{2} \end{matrix}\right.\right], \tag{22}$$

with $A = \prod_{j=1}^{p} \frac{1}{\Gamma(\frac{n-q}{2} - \frac{(j-1)}{2})\Gamma(\frac{q}{2} - \frac{(j-1)}{2})}$.

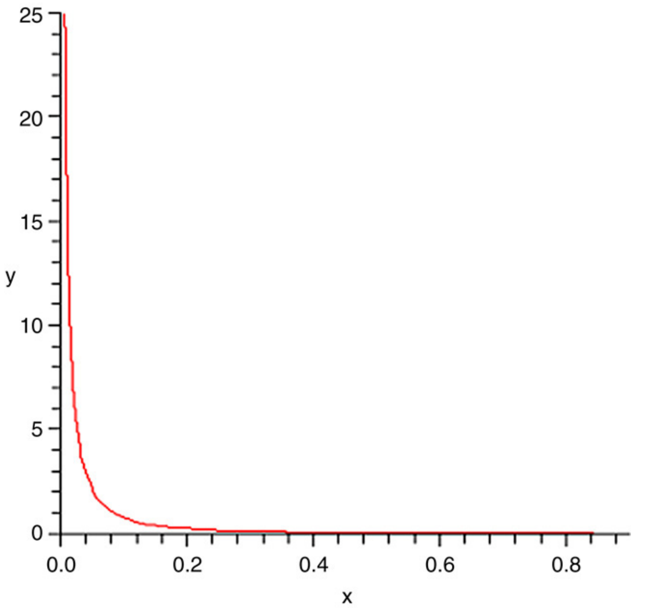

Fig. 2. Density of $\Lambda'(12.5, 4, 8.34)$

In the general case, n, p, q can have non-integral values and we can have $q < p$.

For example, Fig. 2 gives the density of $\Lambda'(n, p, q)$, for $n = 12.5$, $p = 4$ and $q = 8.34$.

Remarks. 1. There are several other statistics closely related to that of Wilks. Let us recall that the Lawley–Hotelling statistic uses the trace of $\frac{\mathbf{A}}{\mathbf{B}}$ and the Roy Maximum root criterion uses its maximum characteristic root, whereas Pillai uses the trace of $\frac{\mathbf{A}}{\mathbf{A}+\mathbf{B}}$.

However, they cannot be treated using entirely the approach given here, and the expressions of some of their densities will be presented in another article.

2. Finally, it should be noticed that, similarly to the univariate case, where the beta prime is also called the gamma–gamma distribution, frequently encountered in Bayesian Statistics [13], $\mathbf{Y} \sim \text{beta}_p^{II}(a, b)$ can also be obtained as the continuous mixture of two Wishart matrix distributions, in the sense of $\mathbf{Y} \sim \mathbf{W_p}(n, \mathbf{W})$, with $\mathbf{W} \sim \mathbf{W_p}(n_0, \mathbf{\Sigma}_0)$.

6. Applications of Wilks's statistics

Wilks's statistic, or a variation of it, has multiple uses in different domains of multivariate analysis, particularly in Multiple Analysis of Variance, in Multiple Regression and in Multiple Discriminant Analysis. In this section we will present the three cases, the likelihood statistic as a rational power of the Wilks's statistic, and the product of two Wilks's statistics and their ratio.

First, as stated in [8], Wilks's statistic could be used to measure the lack of association between two p-vectors $\mathbf{X}$ and $\mathbf{Y}$. If $p \leq q$, and $r_1^2, r_2^2, \ldots, r_p^2$ are the square of the canonical correlations between $\mathbf{X}$ and $\mathbf{Y}$, we then have $\Lambda = \prod_{i=1}^{p}(1 - r_i^2)$, and Λ can be interpreted as a generalization of both the correlation coefficient and the multiple correlation coefficient.

6.1. Likelihood ratios

Let us recall first that Hotelling's T^2 statistic, to test $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ versus $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$, based on a normal sample $\mathbf{X}_1, \ldots, \mathbf{X}_n$, from a population $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is related to $\Lambda(n, p, q)$ by the relation $\Lambda^{2/n} = (1 + \frac{T^2}{n-1})^{-1}$.

(a) *Testing the equality of mean vectors conditional to the equality of covariance matrices.*

Kshirsagar [8, p. 400] considers k independent p-variate normal populations $\mathbf{N}_p(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$, and the 3 null hypotheses:

$\mathrm{H_a}$: $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \cdots = \boldsymbol{\Sigma}_k$
$\mathrm{H_b}$: $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_k$,given that $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \cdots = \boldsymbol{\Sigma}_k$, and
$\mathrm{H_c}$: $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_k$; $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \cdots = \boldsymbol{\Sigma}_k$.

k samples of sizes $n_i, i = 1, \ldots, k$, are available, and let X_i denote the $(p \times n_i)$ matrix of the observations from the ith sample. The matrix of the corrected sums of squares and sums of products of these observations is: $\mathbf{S}_j = \mathbf{X}_j(\mathbf{I} - \frac{1}{n_j}\mathbf{E}_{n_j n_j})\mathbf{X}'_j$, with $\overline{\mathbf{X}_j} = (\mathbf{X}_j \mathbf{E}_{n_j 1})/n_j$ being the vector of sample means and $\mathbf{A} = \sum_{i=1}^{k} \mathbf{S}_i$ is the pooled matrix.

Setting $\mathbf{B} = \sum_{i=1}^{k} n_i(\bar{\mathbf{X}}_i - \bar{\bar{\mathbf{X}}})(\bar{\mathbf{X}}_i - \bar{\bar{\mathbf{X}}})'$, where $\bar{\bar{\mathbf{X}}}$ is the vector of the mean of all sample means, $f_i = n_i - 1, i = 1, \ldots, k$, and $f = \sum_{i=1}^{k} f_i$, $N = \sum_{i=1}^{k} n_i$, we can use $\lambda_b = (\frac{|\mathbf{A}|}{|\mathbf{A}+\mathbf{B}|})^{f/2} = \Lambda^{f/2}$ as a modified likelihood criterion to test $\mathrm{H_b}$. Hence here, we take $\zeta = f/2$ in (19) to obtain the density of λ_b.

For the testing of $\mathrm{H_c}$ above, with its likelihood ratio statistic being equal to the product of the first two in $\mathrm{H_a}$ and $\mathrm{H_b}$, Kshirsagar [8], or Anderson [1, p. 414] can be consulted. For the latter reference, where a product of several factors is considered, with each of them being a product of independent betas, our approach presented in this article and Eq. (11) is particularly suited.

(b) *Testing the independence between k sets of variables*: The likelihood ratio statistic arises under another form in this case, where $\mathbf{X} \sim \mathbf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\mathbf{X}, \boldsymbol{\mu}, \boldsymbol{\Sigma}$ are partitioned as $\mathbf{X}' = (\mathbf{X}'_1, \mathbf{X}'_2, \ldots, \mathbf{X}'_k)$, $\boldsymbol{\mu}' = (\boldsymbol{\mu}'_1, \boldsymbol{\mu}'_2, \ldots, \boldsymbol{\mu}'_k)$ and $\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \cdots & \boldsymbol{\Sigma}_{1k} \\ \boldsymbol{\Sigma}_{21} & & \boldsymbol{\Sigma}_{2k} \\ \cdots & & \cdots \\ \boldsymbol{\Sigma}_{k1} & & \boldsymbol{\Sigma}_{kk} \end{bmatrix}$, where $\mathbf{X}_i$ and $\boldsymbol{\mu}_i$ are $m_i \times 1$ and $\boldsymbol{\Sigma}_{ii}$ are $m_i \times m_i$, $i = 1, \ldots, k$, with $\sum_{i=1}^{k} m_i = p$.

We wish to test the null hypothesis $\mathrm{H_0}$ that the subvectors $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_k$, are independent, i.e. $\mathrm{H_0}$: $\boldsymbol{\Sigma}_{ij} = \mathbf{0}, i, j = 1, \ldots k, i \neq j$.

Let $\overline{\mathbf{X}}$ and $\mathbf{S}$ be the sample mean and covariance matrix from a sample of $N = n + 1$ observations, and let $\overline{\mathbf{X}}$ and $\mathbf{A} = n\mathbf{S}$ be partitioned as:

$$\overline{\mathbf{X}}' = (\overline{\mathbf{X}}'_1, \overline{\mathbf{X}}'_2, \ldots, \overline{\mathbf{X}}'_k) \quad \text{and} \quad \mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \ldots & \mathbf{A}_{1k} \\ \mathbf{A}_{21} & & \mathbf{A}_{2k} \\ \ldots & & \ldots \\ \mathbf{A}_{k1} & & \mathbf{A}_{kk} \end{bmatrix}.$$

Then the likelihood ratio test is used in a left tailed test. Let $\Upsilon = W^{2/N}$. Then when $\mathrm{H_0}$ is true, Υ has the same distribution as $\prod_{i=2}^{k}\prod_{j=1}^{m_i} X_{ij}$, where $X_{ij} \sim \text{beta}(\frac{n+1-(m_i^*-j)}{2}, \frac{m_i^*}{2})$, and are independent, with $m_i^* = \sum_{l=1}^{i-1} m_l$ [12, p. 533]. Using (11), we can now derive the density of Υ under null hypothesis as:

$$h(x) = K\mathbf{G}^{M\,0}_{M\,M} \times \left[x \,\middle|\, \begin{matrix} \frac{n}{2}-1, \frac{n-1}{2}-1, \ldots, \frac{n-(m_2-1)}{2}-1, \ldots, \frac{n}{2}-1, \frac{n-1}{2}-1, \ldots, \frac{n-(m_k-1)}{2}-1 \\ \frac{n^*-(m_2^*+1)}{2}-1, \frac{n^*-(m_2^*+2)}{2}-1, \ldots, \frac{n^*-(m_2^*+m_2)}{2}-1, \ldots, \frac{n^*-(m_k^*+1)}{2}-1, \frac{n^*-(m_k^*+2)}{2}-1, \ldots, \frac{n^*-(m_k^*+m_k)}{2}-1 \end{matrix} \right]$$
$$\text{for } x > 0, \tag{23}$$

with $n^* = n+1$, $M = m_1 + m_2$ and $K = \prod_{i=2}^{k} (\prod_{j=1}^{m_i} \frac{\Gamma(\frac{n+1-j}{2})}{\Gamma(\frac{n+1-(m_i^*+j)}{2})})$.

For small values of the parameters, Anderson [1] has given the expression of this density, as a function of elementary functions, and so did Consul [4] and Mathai [10]. Let us note that, originally, Wilks's statistic, denoted by λ, is related to the same ratio W above, but with $|\mathbf{A}_{ii}|$ being the determinants of the matrix of sample correlation coefficients. Wald and Brookner [20] have studied the distribution of λ, also as a product of independent beta variables, and for a small number of variates, they have expressed the density of λ, and its c.d.f., as functions of λ^k and $\log \lambda$.

6.2. *Product of Wilks's statistics*

The product of independent Wilks's statistics arises in different contexts.

(a) *Partitioning of the E and H matrices according to* $p = p_1 + p_2$: Let the matrix $\mathbf{E}$ be partitioned as: $\mathbf{E} = \begin{array}{c} \begin{bmatrix} \mathbf{E}_{11} & \mathbf{E}_{12} \\ \mathbf{E}_{21} & \mathbf{E}_{22} \end{bmatrix} \end{array}\begin{array}{l} p_1 \\ p_2 \end{array}$ and, similarly, we have: $\mathbf{E_H} = \begin{bmatrix} \mathbf{E_{H11}} & \mathbf{E_{H12}} \\ \mathbf{E_{H21}} & \mathbf{E_{H22}} \end{bmatrix}\begin{array}{l} p_1 \\ p_2 \end{array}$. Let us form the ratio: $U = \frac{|\mathbf{E}|}{|\mathbf{E_H}|} = \frac{|\mathbf{E}_{11}|}{|\mathbf{E_{H11}}|} \frac{|\mathbf{E}_{22} - \mathbf{E}_{21}\mathbf{E}_{11}^{-1}\mathbf{E}_{12}|}{|\mathbf{E_{H22}} - \boldsymbol{E}_{\mathbf{H}21}\mathbf{E}_{\mathbf{H}11}^{-1}\mathbf{E_{H12}}|} = b_{(1)} b_{(2)}$. As shown by Seber [14], $b_{(1)}$ and $b_{(2)}$ are independent, with $b_{(1)} = \Lambda(m_H + m_E, p_1, m_H)$, while $b_{(2)} = \Lambda(m_H + m_E - p_1, p_2, m_H)$.

The density of U can hence be obtained as a product, and we have here:

$\Lambda(m_H + m_E, p_1 + p_2, m_H) = \Lambda(m_H + m_E, p_1, m_H)\Lambda(m_H + m_E - p_1, p_2, m_H)$, which is the same relation as our relation (16). Fig. 3 gives the densities of $X = \Lambda(17, 3, 9)$, $Y = \Lambda(14, 4, 9)$ and of their product $XY = \Lambda(17, 7, 9)$, while the density of their real ratio $Y^* = \frac{XY}{X}$ is distinct from the density of their pseudo-ratio $Y = pr\frac{XY}{X}$.

Remark. In [15, p. 176], the same relation is discussed. However, the notation $\Lambda(\mathbf{x}|\mathbf{y}) = \frac{\Lambda(\mathbf{y},\mathbf{x})}{\Lambda(\mathbf{y})}$ can be misleading, as already discussed, and should be written as $\Lambda(\mathbf{x}|\mathbf{y}) = pr\frac{\Lambda(\mathbf{y},\mathbf{x})}{\Lambda(\mathbf{y})}$.

(b) *Factors of Wilks's statistic*: let us consider $\mathbf{X} = (X_1, X_2, \ldots, X_p)^{\mathrm{T}}$ and $\mathbf{Y}$.

The sequence of hypotheses is [8, p. 320]:

H_0: $\mathbf{X}$ has no relationship with the vector $\mathbf{Y}$
H_1: X_1 has no relationship with the vector $\mathbf{Y}$
H_2: X_2 has no relationship with the vector $\mathbf{Y}$, after elimination of the relationship due to X_1, has no relationship with the vector $\mathbf{Y}$, and in general,
H_k: X_k, $1 \le k \le p$, has no relationship with the vector $\mathbf{Y}$, after elimination of the relationship due to $X_1, \ldots, X_{k-1}$, has no relationship with the vector $\mathbf{Y}$.

We have, $\mathrm{H}_0 = \bigcap_{i=1}^{p} H_i$.

Suppose we want to test X_j, $k+1 \le j \le p$ simultaneously, i.e. $\bigcap_{i=k+1}^{p} H_i$, then, under the null hypothesis that these are true, we have: $\Lambda_{k+1,\ldots,p.12\ldots k}$. However, in this particular case, $\Lambda_{k+1,\ldots,p.12\ldots k}$ can be seen as $\Lambda(n-k, p-k, q)$ [8, p. 325] and the multiplicative relation (16) is again satisfied. Hence $\Lambda_{k+1,\ldots,p.12\ldots k}$ is the pseudo-ratio $pr\frac{\Lambda(n,p,q)}{\Lambda(n,k,q)}$, while the real ratio has its distribution given by (17), and is not a Wilks's statistic.

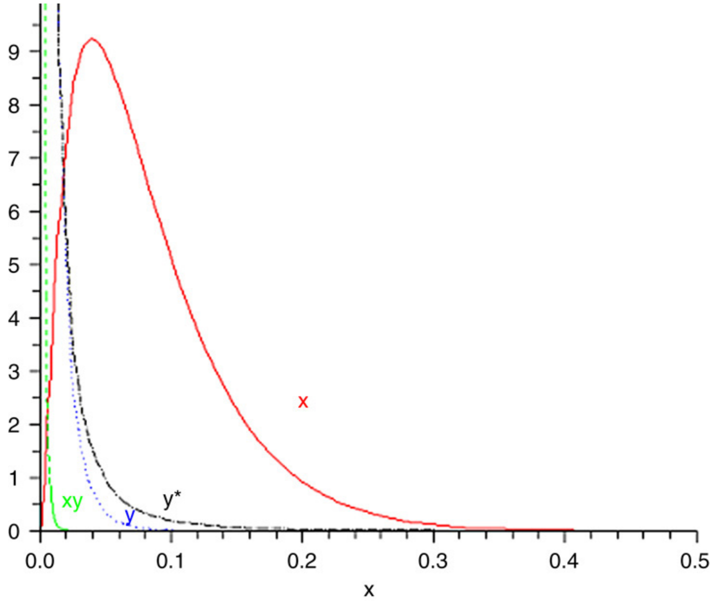

Fig. 3. Densities of $X = \Lambda(17, 3, 9)$, $Y = \Lambda(14, 4, 9)$, their product $XY = \Lambda(17, 7, 9)$, their real ratio $Y^* = \frac{XY}{X}$ and pseudo-ratio $Y = pr\frac{XY}{X}$.

6.3. Ratio of Wilks's statistics

Division into collinearity and direction: In multigroup discrimination, to test the goodness of fit of a single hypothetical discriminant function, we can use the ratio of two Wilks's statistics.

Let us consider the function $\mathbf{h}'\mathbf{X}$ to discriminate among kp-variate normal populations $N_p(\mathbf{X}|\boldsymbol{\mu}_j|\boldsymbol{\Sigma})$, $j = 1, 2 \ldots, k$. As pointed out by Kshirsagar [8, p. 374] the hypothesis of goodness of fit of $\mathbf{h}$ comprises two factors: (a) Collinearity of the means, and (b) direction of the given function.

Bartlett [2] considered first Wilks's lambda $\Lambda(n, p, q) = \frac{|\mathbf{A}|}{|\mathbf{A}+\mathbf{B}|}$ for the relationship between $\mathbf{X}$ and $\mathbf{Y}$, and the relationship of $\mathbf{h}'\mathbf{X}$ with $\mathbf{Y}$, given by $\lambda_A = \frac{\mathbf{h}'\mathbf{A}\mathbf{h}}{\mathbf{h}'(\mathbf{A}+\mathbf{B})\mathbf{h}}$.

Removing the latter gives us the Residual Wilks's Lambda:

$$\Lambda_R = \frac{|\mathbf{A}|}{\lambda_A|\mathbf{A}+\mathbf{B}|} = \frac{|\mathbf{A}|}{|\mathbf{A}+\mathbf{B}|}\Bigg/\frac{\mathbf{h}'\mathbf{A}\mathbf{h}}{\mathbf{h}'(\mathbf{A}+\mathbf{B})\mathbf{h}}.$$

Now Λ_R has the Wilks's $\Lambda(n-1, p-1, q)$ distribution. Hence, we have: $\lambda_A = \frac{\Lambda(n,p,q)}{\Lambda(n-1,p-1,q)}$, whose density is given by (17), which is not a Wilks's statistic.

Fig. 4 gives the density of λ_A for $n = 10$, $p = 4$, $q = 6$.

Factorizing Λ_R again, we have $\Lambda_R = \Lambda_D \Lambda_{(C|D)}$, with $\Lambda_D = \frac{1-\mathbf{h}'\mathbf{B}(\mathbf{A}+\mathbf{B})^{-1}\mathbf{B}\mathbf{h}/\mathbf{h}'\mathbf{B}\mathbf{h}}{\lambda}$ being the direction factor. Hence, the partial collinearity factor $\Lambda_{(C|D)}$ is the pseudo-ratio of two independent Wilks's statistics since it is also a Wilks's statistic, and we have: $\Lambda_R = \Lambda(n-1, p-1, q)$, $\Lambda_D = \Lambda(n-1, 1, p-1)$, $\Lambda_{C|D} = \Lambda(n-2, q-1, p-1)$.

The relation $\Lambda_R = \Lambda_D \Lambda_{(C|D)}$ can be also immediately verified from our relation (16), by an appropriate change of parameters and using equality $\Lambda(n, p, q) = \Lambda(n, q, p)$.

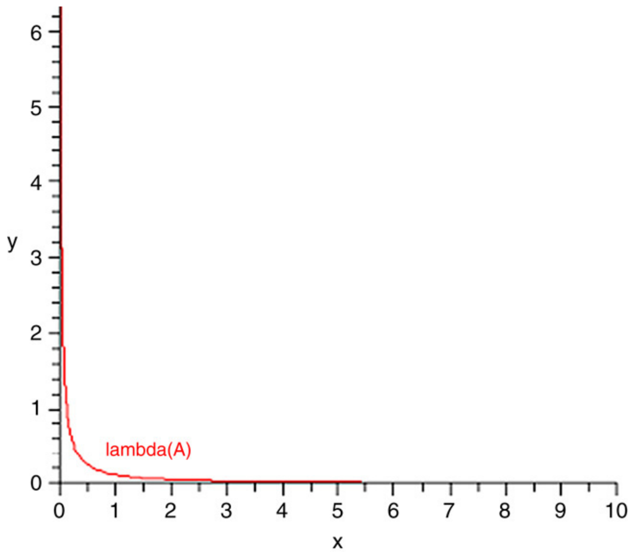

Fig. 4. Density of $\lambda_A = \frac{\Lambda(10,4,6)}{\Lambda(9,3,6)}$.

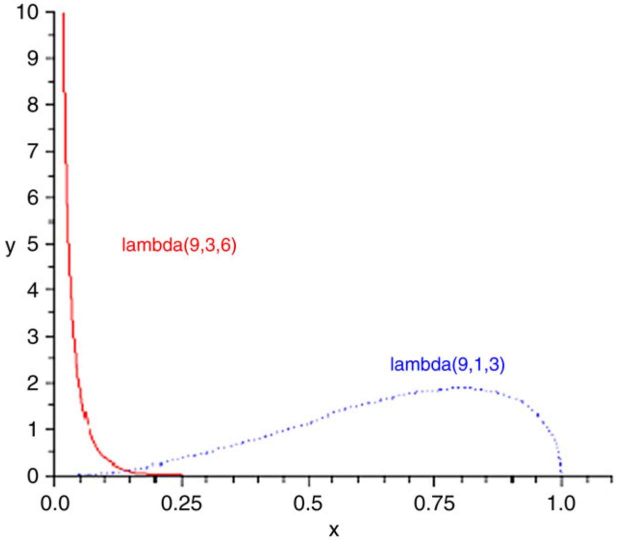

Fig. 5. Densities of $\Lambda(9, 3, 6)$ and $\Lambda(9, 1, 3)$.

Fig. 5 gives the densities of $\Lambda(9, 3, 6)$ and of $\Lambda(9, 1, 3)$, while Fig. 6 gives those of their pseudo-ratio $\Lambda_{C|D} = \Lambda(8, 5, 3)$ and real ratio $\Lambda^* = \Lambda_R/\Lambda_D$, defined on $(0, \infty)$, as given by (17). The two densities do differ from each other.

Similarly, we can consider the collinearity factor Λ_C (called coplanarity factor in [7]) and the partial direction factor $\Lambda_{(D|C)}$, with $\Lambda_R = \Lambda_C \Lambda_{(D|C)}$, where $\Lambda_C = \Lambda(n-1, p-1, q-1)$, $\Lambda_{D|C} = \Lambda(n - q, p - 1, 1)$.

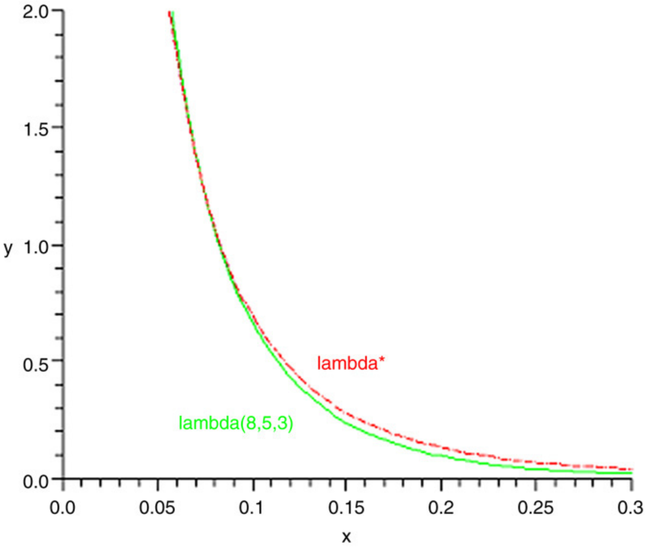

Fig. 6. Densities of pseudo-ratio $\Lambda_{C|D} = \Lambda(8, 5, 3)$ and real ratio $\Lambda^* = \Lambda_R/\Lambda_D$.

7. Conclusion

We have presented the closed form expression of Wilks's statistic, perhaps the most widely used criterion for testing hypotheses in multivariate analysis among the group of selected statistics with similar purposes, such as Pillai trace, the Roy maximum latent root, etc. Also, the availability of the closed formed expressions for the densities of the product and ratio of two independent Wilks's statistics will certainly permit the testing of more complex hypotheses, based on Wilks's statistic.

Acknowledgments

This research was partially supported by NSERC Grant A9249 (Canada). The author wishes to thank Prof. N. Turkkan for several sound advices and Francis Niyomwungere for his excellent support in programming with Maple. Also, the author thanks an anonymous referee for several comments that have helped in improving the presentation of this paper.

References

[1] T.W. Anderson, An Introduction to Multivariate Statistical Analysis, 2nd ed., Wiley, New York, 1984.
[2] M.S. Bartlett, The goodness-of-fit of a single hypothetical discriminant function in the case of several groups, Ann. Eugenics 16 (1951) 199–214.
[3] A. Bekker, J.J.J. Roux, T. Pham-Gia, Operations on the Matrix Variate Beta Type I, Variables and Applications, Unpublished manuscript, University of Pretoria, South Africa, 2005.
[4] P.C. Consul, On the exact distributions of the likelihood ratio criteria for testing independence of sets of variates under the null hypothesis, Ann. Math. Stat. (1967) 1160–1169.
[5] Y. Fujikoshi, V.V. Ulyanov, Error bounds for asymptotic expansions of Wilk's lambda, distribution, J. Multivariate Anal. 97 (2006) 1941–1957.
[6] A.A. Kilbas, M. Saigo, H-Transforms, Theory and Applications, Chapman and Hall/CRC, New York, 2004.
[7] A. Kshirsagar, An alternative derivation of the distribution of the direction and collinearity statistics in discriminant analysis, Calcutta Stat. Soc. 19 (1970) 123–134.

[8] A. Kshirsagar, Multivariate Analysis, Marcel Dekker, New York, 1972.
[9] A.M. Mathai, Saxena, Generalized Hypergeometric Functions with Applications in Statistics and Physical Sciences, in: Lecture Notes, vol. 348, Springer Verlag, New York, 1973.
[10] A.M Mathai, An expansion of Meijer's G-function and the distribution of products of independent beta variates, S. Afr. Statist. J. 5 (1971) 17–90.
[11] A.M Mathai, P.N. Rathie, The exact distribution of Wilks's criterion, Ann. Math. Stat. (1971) 1010–1019.
[12] R., J. Muirhead, Aspects of Multivariate Statistical Theory, Wiley, New York, 1982.
[13] T. Pham-Gia, N. Turkkan, Operations on the generalized F-variables, and applications, Statistics 36 (3) (2002) 195–209.
[14] G.A.F. Seber, Multivariate Observations, Wiley, New York, 1984.
[15] A. Rencher, Multivariate Statistical Inference and Applications, Wiley, New York, 1998.
[16] M. Schatzoff, Exact distributions of Wilks's likelihood ratio criterion, Biometrika 53 (1966) 347–358.
[17] M. Springer, The Algebra of Random Variables, Wiley, New York, 1984.
[18] V.V. Ulyanov, H. Wakaki, Y. Fujikoshi, Berry–Esseen bound for high dimensional approximation of Wilk's lambda distribution, Statist. Probab. Lett. 76 (2006) 1191–1200.
[19] H. Wakaki, Edgeworth expansions of Wilk's lambda statistic, J. Multivariate Anal. 97 (2006) 1958–1964.
[20] A. Wald, R.J. Brookner, On the distribution of Wilks' statistic for testing the independence of several groups of variates, Ann. Math. Stat. (1941) 137–152.